Technical Report 1034

Task Performance in Virtual Environments: Stereoscopic Versus Monoscopic Displays and Head-Coupling

Michael J. Singer, Jennifer Ehrlich, and Stephen Cinq-Mars
U.S. Army Research Institute

Jean-Paul Papin
Centre de Facteurs Humain
Angers, France

December 1995

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

Gene W. Fober
Philip D. Gillis

NOTICES

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1995, December	2. REPORT TYPE Final	3. DATES COVERED January 1993-September 1995
4. TITLE AND SUBTITLE Task Performance in Virtual Environments: Stereoscopic Versus Monoscopic Display and Head-Coupling		5a. CONTRACT OR GRANT NUMBER
		5b. PROGRAM ELEMENT NUMBER 0602785A
6. AUTHOR(S) Michael J. Singer, Jennifer Ehrlich, Stephen Cinq-Mars (ARI), and Jean-Paul Papin (Centre de Facteurs Humain)		5c. PROJECT NUMBER A790
		5d. TASK NUMBER 2111
		5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Instutute for the Behavioral and Social Sciences ATTN: PERI-IF 5001 Eisenhower Avenue Alexandria, VA 22333-5600		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Technical Report 1034

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

The U.S. Army Research Institute for the Behavioral and Social Sciences has an ongoing program of investigation into the requirements for using Virtual Environments (VE) to train dismounted soldiers. As a part of this program, an experiment was conducted investigating the effects of different parameters of VE in the performance of simple, representative tasks. This report provides background information about VE display problems, head-coupling in VE, presence, field dependence, and simulator sickness. The tasks used in the experiment are generic to performance in VEs and would form the basis both of training programs and general soldier tasks.

Visual presentation of the tasks was either through a Stereoscopic Head Mounted Display (HMD) or a Monoscopic HMD, and subjects could either move their Field of View (FOV) by moving their head (coupled) or could not move the FOV (uncoupled). The five tasks used were (1) movement through a sequence of rooms and doorways, (2) acquisition of a fixed target, (3) tracking a moving object, (4) manipulation of an object, and (5) distance estimation. In general, performance of all tasks improved over repeated trials. In the distance estimation task the estimations were relatively worse at shorter distances. However, the error was significantly lessened with stereoscopic presentations, and was also significantly improved when coupling was used, although these factors did not interact

15. SUBJECT TERMS

Dismounted infantry	Stereoscopic displays	Virtual Environments	Head-coupling
Monoscopic displays	Land navigation	Spatial learning	Training

SECURITY CLASSIFICATION OF					
16. REPORT	17. ABSTRACT	18. THIS PAGE	19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
Unclassified	Unclassified	Unclassified	Unclassified	51	

Technical Report 1034

14. Abstract (Continued)

with one another. Performance in the other tasks was not significantly effected by presentation mode or head-coupling.

The distance task errors and the lack of significant differences in performance of the other tasks raise questions on the claimed general gain in task performance through the increased reality of stereoscopic presentations and head-coupling. Theses data indicate that the gains available from stereoscopic presentations or head-coupling are not significantly large for these simple, typical tasks. However, these factors do significantly improve short-range distance estimation.

Technical Report 1034

Task Performance in Virtual Environments: Stereoscopic Versus Monoscopic Displays and Head-Coupling

Michael J. Singer, Jennifer Ehrlich, and Stephen Cinq-Mars
U.S. Army Research Institute

Jean-Paul Papin
Centre de Factors Humain
Angers, France

Simulator Systems Research Unit
Stephen Goldberg, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

December 1995

Army Project Number
2O262785A790

Personnel Systems and
Performance Technology

FOREWORD

The Army has made a substantial commitment to Distributed Interactive Simulation (DIS) and the electronic battlefield for training, concept development, and test and evaluation. The current DIS training system, Simulation Networking (SIMNET), and the next generation system, the Close Combat Tactical Trainer (CCTT), are designed to provide realistic training for soldiers fighting from vehicles, but are not designed to provide training for individual dismounted soldiers. The development of technology that completely replaces real world sensory input with computer-generated simulation, called Virtual Environment (VE) technology, has the potential to provide Individual Combat Simulations (ICS) for the electronic battlefield. However, several research challenges must be overcome before VE technology can be used for practical training applications. One of these challenges is identifying and quantifying the effects of VE system characteristics that influence skill acquisition and transfer.

This report describes the fourth in a series of experiments addressing VE technology for training dismounted soldiers. This experiment was explicitly designed to address the effect of VE system characteristics on skill acquisition. The characteristics addressed were the presentation mode (stereoscopic vs. monoscopic) and visual interaction (head-coupling or not) in terms of their effects on repetitive performance of simple skilled tasks in virtual environments. The simple tasks were designed as components of many different soldier tasks that will eventually be incorporated in VE-based training. The results of the research provides information about the characteristics needed for effective learning and performance in virtual environments.

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) Simulator Systems Research Unit conducts research with the goal of providing information that will improve the effectiveness of training simulators and simulations. The work described here is a part of ARI research task 211, VIRTUE--Virtual Environments for Combat Training and Mission Rehearsal. This experiment was performed as part of a cooperative research and development agreement with the French Ministry of Defense. The experiment was planned with Medecin Chef Des Services Jean-Paul Papin during a visit to the Simulator Systems Research Unit.

ZITA M. SIMUTIS
Deputy Director
(Science and Technology)

EDGAR M. JOHNSON
Director

ACKNOWLEDGMENT

Dr. Cinq-Mars was the primary liaison with the French and co-author of this research plan while a Research Fellow at the Simulator Systems Research Unit in Orlando, FL. He is currently a Human Factors specialist with CAP SESA, Inc. in Paris, France. We also wish to acknowledge the contributions of Research Fellows John Gildea and Dan McDonald. This research could not have been completed without their assistance.

TASK PERFORMANCE IN VIRTUAL ENVIRONMENTS: STEREOSCOPIC VERSUS MONOSCOPIC DISPLAYS AND HEAD COUPLING

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army has committed to using Distributed Interactive Simulation (DIS) and the electronic battlefield for training, concept development, testing, and evaluation. The current DIS training system, Simulation Networking (SIMNET), and the next generation system, the Close combat Tactical Trainer (CCTT), are designed to provide realistic training for soldiers fighting from vehicles, but are not designed to provide the same benefit for individual dismounted soldiers. Virtual Environment (VE) technology, which replaces real world sensory input (e.g., visual displays) with computer-generated simulation, has the potential to provide Individual Combat Simulations (ICS) for training, concept development, testing, and evaluation on the electronic battlefield. Two key distinguishing characteristics of VE systems are the complete replacement of visual stimuli with a visual display and the use of head-coupling to synchronize changes in the visual display with head movement. The presentation of complex stereoscopic visual displays synchronized to head movement is a major computational problem, and together with the computing equipment, is a major cost driver in ICS. Research is needed to identify and quantify the effects of these and other VE systems characteristics that may influence skill acquisition of typical psychomotor activities within ICS.

Procedure:

The central VE factors in this experiment are the effects of different display modes and the use of head-coupling. These VE characteristics must be evaluated in the context of typical Army tasks or activities. Previous research at the Simulator Systems Research Unit developed a battery of tasks for performance in VEs that cover a range of psychomotor activities which are characteristic of tasks performed by Armed Forces personnel to investigate different VE characteristics. In this experiment, a subset of these representative tasks was performed using stereoscopic or monoscopic presentations in a helmet-mounted-display, and head-coupled or fixed-scene tracking. The activities used were control of self-movement in restricted areas (corridors), object manipulation, tracking to fixed or moving targets, and estimating distances to a familiar object. This experiment also considered the acquisition of skill through repeated trials. Forty-eight subjects participated in the experiment, and a variety of background tests were administered

to ensure a subject pool with good vision, no prior experience with virtual environments, and no history of motion sickness. Gender and Field Dependence (as measured by the Hidden Figures Test) were used to assign subjects to conditions.

Findings:

As expected, all subjects improved in the performance of the tasks during the repeated trials. There were significant differences found with the head-coupling and the viewing conditions only in performance of the Distance Estimation task. In that task, near distances were more correctly estimated with head-coupling than with no head-coupling, with little difference at greater distances. This was also the case with the viewing condition, where stereoscopic presentation improved the estimation of close distances over monoscopic presentation, while greater distances were similar in accuracy. In addition, the improvement in distance estimation differed over trials, with shorter distances improving while longer distances did not. In all cases the relative accuracy at longer distances (20 and 30 feet) was much better than at shorter distances (2.5, 5, and 10 feet). One of the Simulator Sickness Questionnaire subscale differences (pre-administration to post-administration) revealed significantly greater nausea associated with the stereoscopic presentation.

Utilization of Findings:

This research provides information for both application and further research. Individual Combatant Simulations (ICS) can probably use monoscopic displays for certain kinds of tasks, easing the processing requirements in the simulation. The additional processing and display equipment needed for stereoscopic displays should only be considered for tasks requiring relatively short-distance estimations or visually close work. Research in VE configurations should focus on the complex interaction of visual and task characteristics, the human factors aspects associated with stereoscopic presentations and head-coupling. Research should also be conducted into the quality and accuracy of presentations at short distances.

TASK PERFORMANCE IN VIRTUAL ENVIRONMENTS: STEREOSCOPIC VERSUS MONOSCOPIC DISPLAYS AND HEAD COUPLING

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

TASK PERFORMANCE IN VIRTUAL ENVIRONMENTS: STEREOSCOPIC VERSUS MONOSCOPIC DISPLAYS AND HEAD-COUPLING

In a downsized Army, training time and cost effectiveness are of vital importance. The Army has made a substantial commitment to the use of Distributed Interactive Simulations (DIS) for training and testing soldiers, developing and testing operating doctrine, and evaluating concepts (see Knerr, Goldberg, Lampton, Witmer, Bliss, Moshell, & Blau, 1994; Jacobs, Crooks, Crooks, Colburn, Fraser, Gorman, Madden, Furness, & Tice, 1993). The goal of DIS is to decrease the high cost of operational rehearsal and training, increase the testing of new doctrine, and enable the evaluation of new equipment concepts with minimal cost. Historically, much of the networked simulation has focused on vehicles, as a natural outgrowth of the individual weapon and vehicle simulators the military has been building for decades. As a result of that focus, and because of the difficulty of even minimal simulation, the dismounted infantryman has been left out of DIS. New efforts (e.g. 21st Century Land Warrior) are emphasizing the importance of individual soldiers, and indicate the need for Individual Combatant Simulation (ICS).

Virtual reality (VR) or virtual environments (VE) technology presents a new type of simulation technology that specifically focuses on the individual. A VE system fully occupies one or more of an individual's perceptual systems in an alternate world or environment through wearable equipment such as head-mounted displays and position trackers. This technology is now being investigated for potential use in providing the dismounted infantryman with Individual Combatant Simulation (ICS). The U.S. Army Research Institute, Simulator Systems Research Unit, research program is designed to investigate the characteristics of VE systems for training dismounted soldiers (Knerr, et al., 1994).

The performance, psychological, and physiological effects of VE use require detailed study if VEs are to be used for training dismounted soldiers. One area of investigation involves the requirements of visual display systems, such as Helmet-Mounted Displays (HMD). Humans normally perceive stereoscopically, in three dimensions, through the merging of two slightly offset views of the world (one from each eye). An HMD can preserve some of the stereoscopic nature of normal vision by providing fixed-offset views (one to each eye). This is done by computing an entire visual display and selecting two different but overlapping areas for display through the two screens in the HMD. Providing stereoscopic views requires increased computation and more complex control than monoscopic views (which provide identical images to each eye). Stereoscopic presentation therefore increases the cost and complexity of individual point of view simulations. However, some researchers believe that the presentation of stereoscopic, fixed-offset views can cause side

effects and physiological stress (Rushton & Wann, 1993; Wann, 1993). In addition, many of the cues used for depth perception are present and usable in traditional monoscopic presentations (such as flat screens or large dome projections), and have been successfully used in simulations and training devices for decades. The research question is whether there are training or practice conditions that require stereoscopic presentations for acceptable performance in VE.

In order to frame the conducted experiment appropriately, this report first provides an overview of both the physiology and optics underlying vision, with a focus on issues relevant to stereopsis and visual processes in virtual environments. This discussion provides background for an experiment designed to evaluate whether or not performance gains in HMDs occur under stereoscopic or monoscopic conditions. The experiment also investigates several other important VR issues, such as the influence of head-coupled visual displays on performance, simulator sickness problems, and the perception of presence. Finally, the research also addresses possible influences of cognitive style differences (measured by the ability to distinguish elements within a complex pattern) on performance.

Vision and Helmet-Mounted Displays

Stereoscopic vision (stereopsis) provides an important cue for the perception of depth in the visual field. This cue is the disparity in an objects projected position on the visual fields in each eye. A number of visual processes and phenomena work together to enable stereoscopic vision. First, the eyes must coordinate their movements to fixate on the object of regard, placing the image on the central portion of each retina. This process is referred to as vergence. In order to focus on objects, the shape of the lens is changed through the action of muscles attached to the capsule of the lens in three distinct rings. This is referred to as accommodation. The amount and origin of bending is unique for every focal distance (Koretz & Handelmann, 1988). Vergence and accommodation are considered to be weak nonvisual cues for stereopsis (Murch, 1973). Finally, the images from both eyes must be fused together by processing in the brain to form one percept. At this point many other perceptual cues are derived from the mix to support the correct experience of three dimensions, the perception of depth. Many of these cues are classed as monoscopic, since they do not require two eyes for acquisition. Stereopsis is a special cue because it occurs even in the absence of monocular cues, as long as the patterns of adjacent points of brightness are the same in the visual field of each eye (Murch, 1973).

As mentioned above, both eyes must move together to fixate on the object. However, they can move simultaneously in a number of ways. When both eyes move in the same direction, it is called

conjunctive (or version) eye movements. If the eyes move in opposite directions, it is called disjunctive (or vergence) movements. As an observer looks from a distant object to a closer one, both eyes need to turn inward. When focusing from near to far, both eyes must rotate outward. Technically then, **convergence** is the vergence movement of turning of the eyes inward to focus on a close object and **divergence** would be the vergence movement of turning the eyes outward to focus from a near object to a distant one. Regardless of the type of movement, the movements of the two eyes are simultaneous and coordinated. In other words, the eyes move as a unit -- as if they were one eye instead of two (Ono, 1991).

Many other monoscopic cues are added to the stereoscopic cues during processing, supporting the perception of depth and distance in the environment. Among these are relative size, interposition, linear perspective, texture gradient, motion parallax, light orientation and shading, as well as contrast, clarity, and brightness (Murch, 1973). Many of these cues are interpretive, with the perception of depth based on assumptions about the size, interposition, and spacing of familiar objects. Relative size refers to the perceiver's assumption that certain objects are the same size, and that the difference in visual angle subtended is used to perceive differences in distance. Interposition refers to the assumption that an object which is partially obscuring another means that the covering object is closer. Linear perspective is commonly used in paintings, where objects that are presumed to be parallel or repeated arrangements with consistent spacing appear to converge with greater distance from the viewer. Texture gradients are most apparent with regular surfaces, and arise from a decrease of surface element detail in a visual area being perceived at increased distances. Motion parallax arises from observer movement that effects the relative position and apparent movement of environmental objects. Objects closer than the moving observer's fixation point appear to be moving backward while objects further away from the fixation point appear to be moving forward. Light orientation and shading refers to the assumed orientation of the sun, which allows inferences about relative distances of objects. Contrast, clarity, and brightness provide indications of distance because in the real world these characteristics decrease as distance increases.

In most cases, humans do not experience the complex merging of monoscopic and stereoscopic information. The process happens automatically, and humans perceive as if with only one eye, and as if that eye were located midway between the anatomically spaced eyes (Ono, 1991). The stereoscopic disparity cue is functional only for a rather limited range, at maximum of ten to twenty meters from the observer (Rinalducci, 1993), and is strongest at very short distances. Beyond this range, the axes of the eyes are aligned at what is essentially optical infinity

for our visual system, and the views from the two eyes are no longer different enough to produce disparity-based three-dimensional cues. At those distances, for similar reasons, no further changes in vergence or accommodation can occur. In such cases, only the monocular depth cues mentioned above can be used by the observer.

In traditional simulation environments, an image is presented on a stationary, flat-screen visual display. These displays produce the illusion of depth through the use of common monocular depth cues (e.g., linear perspective, occlusion, parallax, etc). In many VR systems two small displays are mounted into a headset, referred to as a Helmet-Mounted Display (HMD). These screens are mounted in front of the user's eyes, providing visual stimulation to the user. Some of these HMDs attempt to provide stereoscopic vision by presenting appropriately disparate views to each eye on the separated screens. The paired displays attempt to mimic the disparate views of objects normally received by the eyes due to their physical separation in the head (referred to as binocular disparity). When these separated views are merged in perception, wearers can experience three dimensional stereoscopic depth. Theoretically, this HMD configuration should be able to present the presumably important additional depth information that is normally gained through binocular disparity.

Vergence movements and accommodation are normally linked together (Robinett & Rolland, 1992; Lipton, 1991). However, in an HMD, the parallax or inter-pupillary distance is fixed at the distance between the display screens within the HMD. The vergence angle is then adjusted by the software to compensate for the distance from the eye to the HMD lenses and the disparity between the images sent to the two display screens. At the same time the focal (accommodation) distance is fixed at the location of the display screens within the helmet. Therefore, accommodation and vergence may be de-coupled when viewing objects presented at different distances (differing positions in the computer generated displays) in an HMD (Robinett & Rolland, 1992). This rendering, therefore, can produce a non-normal state of focal and vergence adjustment. This is the opinion of Patterson and Martin (1992):

> [In a] situation where eye movements are permitted and an operator fixates on a disparate stimulus appearing in a depth plane different from that of the display screen, the stimulus for accommodation (display screen) may be at one distance while vergence angle is appropriate for another distance (disparate stimulus), thereby producing a mismatch between accommodation and vergence. Such situations are known to produce much discomfort for the operator. (p. 675)

Thus, the HMDs by their very nature may produce sickness symptoms, which is discussed later in the introduction.

These opinions are supported by research that suggests that the visual system may not be able to comfortably maintain such a split, and visual stress will result (Wann, 1993). In fact, evidence of such physiological stress has been found after merely 10 minutes in an HMD (Rushton & Wann, 1993). Rushton and Wann found evidence of binocular stress in distance vision, binocular fusion, and convergence processes in a task requiring transitions between near and far focusing (biking). In addition, over half of their subjects evidenced visual stress symptoms such as blurred vision. Their reasoning for this stress is that in the closed HMD system, there are no respites, no actual 3-dimensional objects at different distances with matching focal and vergence values to attend to even briefly.

The research reviewed has focused on the processes involved in normal stereoscopic vision, and how this may be affected by HMD systems. However, monoscopic displays are also possible in these HMD systems. In other words, the systems can present the same view to both eyes, and adjust the convergence angle via the software, to produce a fused image without the disparity-based stereoscopic cue. This possibility is often overlooked because of the allure of having a stereoscopic display and the assumption that the stereoscopic display is superior because it is closer to the real world. Yet, it is not clear that there are any performance gains in stereoscopic as opposed to monoscopic displays rendered in this way.

Several studies which did compare one form of stereoscopic and monoscopic presentations found stereoscopic viewing to be superior for depth perception over monoscopic conditions (Yeh & Silverstein, 1992; Sollenberger & Milgram, 1993, using a three-dimensional path tracing task). However, these studies used goggles with shutters that alternately opened and closed over each eye to produce stereoscopic vision (offset views presented to each eye in turn by a standard cathode ray tubes [CRT]). The shutters moved too quickly to be consciously detected, but did produce the perception of depth. The monoscopic conditions in these studies were simply flat CRT computer screens. It is not clear whether or not similar results would be produced with stereoscopic and monoscopic views produced using HMD display equipment.

The research reported here addresses the relative performance differences in a stereoscopic vs. monoscopic HMD using simple tasks. These simple tasks are drawn from the Virtual Environment Performance Assessment Battery (VEPAB) developed by Lampton, Knerr, Goldberg, Bliss, Moshell and Blau (1994). The VEPAB tasks are representative of tasks and activities that would be performed in VE during training,

rehearsal, or performance measurement activities. As stereoscopic viewing has been shown to be superior with at least one equipment configuration and a highly figural task (path tracing, Yeh & Silverstein, 1992), a more common configuration and representative task set are addressed in this research with the expectation that stereoscopic presentations will provide the basis for improved performance.

Head-Coupling

As described before, an HMD positions one or more screens in front of each eye so that during head movements the view remains located directly in front of the eyes. Through the use of head-coupled tracking devices, such as a Polhemous™ sensor, the orientation and position of the user's head is monitored and can be recorded. This information can also be fed back to the computer, so that views corresponding to head position can be generated for presentation as the head moves around.

The same research indicating that binocular disparity may not be the most effective source of depth information in VR also indicates that a more important factor may be head-coupled interaction with the display (Sollenberger & Milgram, 1993; Ware, Arthur, & Booth, 1993). Additionally, these studies suggest that a dynamic rather than static display is key in the perception of depth. Sollenberger and Milgram (1993) conclude that the kinetic depth effect (the perception of depth when a two-dimensional object undergoes three-dimensional rotation around an axis) provides more effective depth information than straight-forward, non-rotational stereoscopic presentation. In a stereoscopic display, if the viewer can observe a static stimulus from multiple angles, performance, though not as good, almost equals that of a stereoscopically presented rotating stimulus. However, the research also indicates that if the stimulus is presented monoscopically, movement is necessary. Viewing multiple angles in a monoscopic display is equivalent to a simple two-dimensional display. The VR/HMD system, rather than presenting a single view with rotation via a stationary screen, allows the individual to move their point of view around in the environment. Although the stimulus is viewed from different angles, these perspectives are generated not by the movement of the stimulus, but by the movement of the viewpoint.

Sollenberger and Milgram (1992) caution that although the ability to rotate a display may be important for visual tasks requiring physical connections between objects and depth discrimination, stereoscopic information may be more vital than rotational information when relative spatial locations, viewing direction, manipulation, and navigation are involved. Support for this caution also comes from research in spatial knowledge acquisition which found no difference between head-tracked and

joystick-manipulated viewing in learning a route through a building (Bailey, 1994).

Although a dynamically head-coupled visual presentation is a further step towards realism in simulations, it is not without problems. The extra computational time required for dynamic head-tracking may introduce disturbing lag times in the computer's screen updating capabilities. There is evidence that a dynamic head-tracked display may increase the incidence of simulator sickness symptoms (Bailey, 1994; Biocca, 1992).

The research experiment reported here crossed dynamic head-tracking or an alternative fixed view presentation with the stereoscopic or monoscopic viewing conditions in the task environments. This will provide information about the usefulness of head-tracking in the performance of simple tasks, furnish additional evidence about the incidence of simulator sickness symptoms associated with dynamic head-tracking, and may also provide information about the interaction of viewing conditions and dynamic head-tracking.

Cognitive Style

In addition to the effects that perceptual mechanisms may have on performance in a virtual environment HMD, there are many different individual characteristics that can subtly affect one's performance of different types of tasks. One possibility is that perceptual style or ability can enhance or retard certain types of skill acquisition.

Field Independence refers to a skill dimension or range of abilities that enable an individual to distinguish coherency in complex perceptual fields. Someone who is Field Independent is more able to separate and distinguish complex elements within a more complex stimulus field (Buckin, 1971). The Field Independent person can more easily separate an item from the context and break up or restructure an already organized perceptual field. Finally, this type of person apparently deals actively with the environment, and is concerned with the details of the environment. In contrast, the Field Dependent person is more global and passive, less able or inclined to re-organize the visual field, simply accepting what is given or first perceived.

Given this framework, a Field Independent person may be able to perceive depth from two dimensional cues more effectively than a Field Dependent person. Consequently, the Field Independent person may not need stereoscopic presentation, and may generally perform better than the Field Dependent person. If this hypothesis is true, it would have implications for equipment design or personnel selection. Rather than investigate this complex of characteristics as a controlling variable, in this

research it will be investigated as a possible intervening variable or confound.

Simulator Sickness

Simulator sickness is a common and constant problem with simulators and training devices. Simulator sickness is similar in effects and outcomes to motion sickness. The symptoms of simulator sickness do resemble those of motion sickness; e.g., nausea, headache, stomach awareness, disorientation, sweating, vomiting, fatigue, eyestrain, etc. There seem to be important differences between motion sickness and simulator sickness (Kolasinski, 1995). The primary difference is that sickness often occurs in simulators when the operator is not moving (as is always the case with motion sickness) and there are a multitude of simulator characteristics involved. Motion and simulator sickness may be different phenomena with potentially different origins (Kolasinski, 1995), but whatever the cause, performance decrements may result. For example, Bailey (1994) found that higher post-VE sickness scores correlated with poorer learning of a route through a building. Since VE will be used in many different simulation and training applications, investigation of possible causes must begin now. Although a thorough review of the complex realm of simulator sickness is beyond the scope of this paper, several relevant issues are important and will be discussed here in order to clarify our measurement of the phenomenon in this research (see Kolasinski, 1995 for a good review).

The most widely regarded theory on the origin of motion sickness is the cue conflict theory (discussed in Kolasinski, 1995). The cue conflict theory proposes that sickness is the result of discordant information from the visual and the vestibular systems. In other words, the visual system may be registering motion based on the images presented on a screen, while the vestibular system senses some incongruent or no actual motion. To the extent that one is unable to adequately rectify this disparate information, simulator sickness results. For example, the feeling of vection (illusory self-movement) seems to be a key factor in producing simulator sickness (Hettinger, Berbaum, Kennedy, Dunlap, & Nolan, 1990). Limitations in the visual display may also produce simulator sickness because the visuals are providing non-normal information. As mentioned above, the addition of head-coupled visual tracking may increase the likelihood of simulator sickness through increasing flicker, and the attendant slow refresh or update rates may also cause problems (Kolasinski, 1995). Kennedy, Lilienthal, Berbaum, Baltzley, & McCauley (1989, cited in Kolasinski, 1995) pointed out that a greater Field Of View (FOV), as provided in the HMD, may also increase the likelihood of simulator sickness by increasing the sense of vection associated with the peripheral field.

In our research, we are interested in the differential effects of monoscopic and stereoscopic displays on the incidence or severity of simulator sickness. Testing the effects of experimental conditions requires a reliable and valid measure of simulator sickness.

Simulator Sickness Questionnaire

Kennedy, Lane, Berbaum, and Lilienthal (1993) have designed a simulator sickness questionnaire in an attempt to delineate symptoms related to simulator exposure. They identified self-report scaled symptoms associated with simulator sickness, and derived three factor analytic subscales in addition to a **Total Severity** measure based on these items. The first is the **Nausea** scale, composed of general discomfort, increased salivation, sweating, nausea, difficulty concentrating, stomach awareness, and burping symptoms. Second is the **Oculomotor Discomfort** scale, with items addressing general discomfort, fatigue, headache, eye strain, difficulty focusing, difficulty concentrating, and blurred vision. Finally, there is the **Disorientation** scale, with symptoms covering difficulty focusing, nausea, fullness of head, blurred vision, dizziness with eyes open, dizziness with eyes closed, and vertigo. The symptom scores are summed and then weighted to arrive at the subscale values, and the total score is derived in a similar fashion.

In this experiment, no great degree of simulator sickness is anticipated, however, the test is being used in every experiment conducted in our research program (Knerr, Goldberg, Lampton, Witmer, Bliss, Moshell, & Blau, 1993; Witmer, Bailey, & Knerr, 1995) and was administered in this experiment in order to adequately evaluate any effects that might occur. Further, data collected in this experiment will eventually be combined with data from other experiments for various analyses investigating possible relationships with VE characteristics, perceptions of the presence in the immersion experiences, and task performance. That work will be reported in a separate report.

Postural Stability

An alternate theory on the possible basis of simulator sickness has been proposed by Riccio and Stoffregen (1991). Unlike the cue conflict theory, theirs is an attempt at a more predictive model of simulator sickness. They "... hypothesize that sickness results when the individual lacks or has not yet learned strategies for maintaining postural stability." (reviewed in Kolasinski, 1995). In the course of our groups research, a method for measuring postural stability has been developed. The subject is required to balance on one foot, with the other lifted while arms are crossed at the chest, eyes are closed, and a blank HMD is worn. The individual is allowed to stand on either leg. The measurements are the magnitude of movement in the X (front-

to-back) and Y (side-to-side) axes taken from the head during this process.

It is possible that the cue conflict and postural stability theories are not necessarily mutually exclusive. For example, disorientation and fatigue are symptoms of simulator sickness, which may also influence one's ability to maintain a stable posture. However, Kennedy, Fowlkes, and Lilienthal (1993) found simulator sickness occurrence without loss of muscular coordination and concluded that the processes producing each are different.

In our research, we measure postural stability in conjunction with the simulator sickness and other questionnaires. If there is a relationship between the postural stability and common components of simulator sickness, this may provide information on the nature of simulator sickness as generated in VR situations. By measuring postural stability both before and after the experimental conditions, an understanding of the relationship between muscular-based stability, simulator sickness indications, and VR conditions may be developed that provides insight to the origins of simulator sickness. These analyses are ongoing, and will not be addressed in the analyses for this experiment. The data collected for this experiment will be reported in terms of effects and relationships for this experiment only. We expect that there will be no differences between subjects assigned to the different conditions in the experiment. Finding differences in the post-experiment test would be surprising, and should correlate to the Simulator Sickness Questionnaire. However, no direct relationship is expected between the postural stability measure and performance of the experimental tasks.

Presence

Presence is the subjective feeling of being immersed in one environment, while actually being situated physically in another. A number of factors potentially related to this feeling of presence have been delineated (see Witmer & Singer, 1994 for review). These factors include Control (degree, immediacy, anticipation, mode, and physical/environmental modifiability); Sensory (sensory modality, environmental richness, multi modal presentation, consistency of multi modal information, degree of movement perception, and active search); Distraction (isolation, selective attention, and interface awareness), and Realism (scene realism, consistency of information with objective world, meaningfulness of experience, and separation anxiety/ disorientation).

Based on these theoretical factors, Witmer and Singer (1994) devised and tested several versions of an **Immersive Tendencies Questionnaire (ITQ)** and a **Presence Questionnaire (PQ)**. Based

upon reliability, correlational, and cluster analytic studies on several versions of the questionnaires (Witmer & Singer, 1994; Singer, Witmer, & Bailey, 1994), several subscales were identified that have a negative relationship with Simulator Sickness. The **Involvement** subscale consists of items on the propensity to actively get involved while witnessing some activity. The **Focus** subscale measures current mental alertness and concentrative skills. The **Games** subscale addresses the extent of experience with arcade and video games. Three subscales were also developed from the items in the **Presence Questionnaire**. The largest of these (in terms of the number of items) is the **Involved/Control** subscale, which deals with the subjective experience of control over and involvement in events in the VR. The **Natural** subscale addresses how natural the experience felt and how consistent with reality the VE experience was. The **Interface** subscale indicates possible display or control interference or distraction during the VE experience.

Witmer and Singer (1994; Singer, Witmer, & Bailey, 1994), also discovered evidence for an inverse relationship between simulator sickness and the Presence measure. They cite three possible explanations for this finding. First, factors associated with increased presence may decrease or prevent simulator sickness. Second, when simulator sickness symptoms are felt, one's attention turns to them, rather than the VR and moving through the experience. This obviously would limit the attentional resources required for coping with the VR experience. Finally, there may be some factor, or set of factors, that mediate both presence and sickness.

We anticipate that there will be some relationship between scores on the Immersive Tendencies Questionnaire, the Presence Questionnaire, and performance of at least one or more of the tasks. We expect that the subjective experience of presence as measured by the Presence Questionnaire will be greater in the more naturalistic conditions (Stereoscopic presentation with head-coupling), and lower when in the least naturalistic condition (Monoscopic presentation without head-coupling).

Research Objectives

The experiment was designed to investigate the VR effects of display characteristics and head-tracking for training in distributed interactive systems. This experiment investigated the relative performance effectiveness of stereoscopic vs. monoscopic displays, and head-tracking vs. no head-tracking in the performance of simple, generic tasks. The tasks used were distance estimation, self-movement, object manipulation, and two different kinds of tracking (Lampton, Knerr Goldberg, Bliss, Moshel, & Blau, 1994). These tasks cover a range of psychomotor activities that share characteristics with tasks performed by armed forces personnel. For example, soldiers may be required to

enter a building and move through corridors (self-movement), manipulate objects (object manipulation), track targets (tracking & targeting tasks), and estimate distances to familiar objects (distance estimation). The first hypothesis of interest is on the effect of stereoscopic and monoscopic displays; the null hypothesis being that there will be equivalent performance using the two types of displays. In addition, the effects of head tracking will be examined; with the null hypothesis being that there is no difference between head-tracking and non-head-tracked performances. If these hypotheses are rejected, then we will have evidence for the common claims that stereoscopic head-tracked interaction is more effective than monoscopic and uncoupled interaction. During the experiment Simulator Sickness, Presence, and Field Dependence will also be investigated.

Method

Subjects

Thirty-six males and twelve females were recruited for this study. Subjects ranged in ages from 18 to 50 years with a mean of 23.6 years. They were paid $5.00 per hour or given course credit for their participation. All subjects had normal, or corrected-to-normal vision (20/33 for near point acuity), no color vision deficiencies, and good stereopsis. Vision tests included a Snellen eye chart, a near point acuity test, the Ishihara color test, a stereopsis test, and lateral and vertical phoria tests. Subjects failing any one of these tests were excluded from the experiment. Subjects with prior experience in research in virtual environments were also excluded, as were subjects with a history of seizures or motion sickness.

Materials and Equipment

The Virtual Environment tasks were developed with Sense8 software, and are a subset of the environments previously used by Knerr, Goldberg, Lampton, Witmer, Bliss, Moshell, & Blau (1993). Five generic tasks were used, as follows:

Doorways is an interior movement task requiring controlled movement through ten rooms with doors at various locations on the opposing walls (no doors are in direct line). Performance measures include time to cross each room and the number of collisions (with walls or doorways) in each room.

Bins is a micro-manipulation task that presents a square of nine bins, with one bin marked by an X as the goal, a target ball in one bin, and the three dimensional, crossed-line cursor starting in the center bin. The **Bins** task requires the subject to maneuver the cursor within the grid to "grab" a target ball by

placing the cursor inside the ball and pressing the button on top of the joystick. The subject then moves the target ball into a different target box using the joystick. Performance measures include time to "grab", total performance time, and accuracy (measured by the number of grabs made and the successful completion of the task within a 45 second time limit).

Fixed Tracking requires the subject to place the three dimensional, crossed-line cursor on a stationary ball-shaped target. The target, a 0.7 ft diameter ball, appears in the room between 5 and 19.5 ft away, at different fixed locations in the three dimensional room on each trial. The target changes color when the cursor is on or within the target. The target disappears after approximately 20 seconds elapses. The performance measure is the percentage of the total trial time during which the cursor is kept on the target and the time to first intercept.

Moving Tracking requires the subject to use the three dimensional, crossed-line cursor to track a moving ball. The ball originates at one plane of the room (walls, floor, or ceiling) and moves across the viewing area in a straight line with a randomly generated slope, and disappears upon reaching another surface. The path can be in any direction and at any angle to the viewpoint, for instance, the ball could traverse from the lower-right near corner to the upper-left far corner or vice versa. The target takes between thirteen (13) and nineteen (19) seconds to traverse the room. The target changes color when the cursor is within the target radius. The performance measure is the percentage of the total trial time during which the cursor is kept on the target and the time to first intercept.

Distance Estimation is a direct distance estimation task, requiring the subject to identify an object (a soldier) judge its height, and estimate its distance as it moves toward the viewer. The object begins at 40 feet from the observer, and the experimenter records when the subject reports that the soldier has moved to the target distances: 30 ft, 20 ft, 10 ft, 5 ft, and 2.5 feet away from the subject's position. The performance measure is the accuracy of distance judgements. There was no overt feedback about the accuracy of judgment during or after the three trials.

A Virtual Research Flight Helmet[tm] with 83 degree field of view (FOV) was chosen as the visual display for the experiment. The Flight Helmet has a resolution of 360 x 240 color pixels in

each lens. The Flight Helmet uses LCD technology with LEEP™ optics which align the two lenses. Stereoscopic images were generated by a Silicon Graphics Reality Engine, which sends slightly different images to each lens. In order to achieve monoscopic presentation, the same view was projected by both lenses and Fresnell lenses were used to help merge the identical images into one. A Polhemous IsoTrack provided head tracking. Finally, subjects used a six-degree-of-freedom joystick to control their interaction (movement, tracking, or manipulating) in the virtual world.

Questionnaires and test materials included tests of vision, information processing style, simulator sickness measures, and evaluations of immersive tendencies and experienced presence in the VE. Vision tests included a Snellen eye chart, a near point acuity test, the Ishihara color test, a stereopsis test, and lateral and vertical phoria tests. The Hidden Figures test (Ekstrom, French, Harman, & Dermen, 1976) was used to measure the field dependence level of the subjects. The Immersive Tendencies Questionnaire and Presence Questionnaire (Witmer & Singer, 1994; Singer & Witmer, 1994) were used to assess concentrative or focusing tendencies and the subjective experience while performing tasks in the VE, respectively. As mentioned above, the Simulator Sickness Questionnaire (Kennedy, Lane, et al, 1993) was used to assess changes in health over the course of the experiment and a postural stability test was conducted both before and after the VE experience. A general VR effectiveness questionnaire was developed and administered to assess the quality of stimulus presentation during the experiment.

Procedure

As introduced above, there were two main between subjects factors in the experiment: visual presentation (stereoscopic or monoscopic view) and coupling (dynamic head-tracking or fixed view). There were repeated measures on each experimental task -- 30 each for the **Bins**, **Fixed Tracking**, **Moving Tracking**, and **Doorways** task, as well as 3 estimates at each of 5 distances on the **Distance Estimation** task.

Subjects were assigned to one of the four experimental groups, within a balancing arrangement for gender and field dependence. Subject assignment to the different conditions was based on field dependence as measured by the Hidden Figures Test in order to balance any effects in the experimental conditions. Three field independent and nine field dependent subjects participated in each group, with the exception of the stereoscopic-non-head-tracked group, in which there were four field independent and eight field dependent subjects. Each group had three women and nine men. Gender was crossed with field dependence in the experimental groups.

Subjects completed an Informed Consent Form after receiving an overview of the experiment and being informed of possible side effects of participation, e.g., simulator sickness. Subjects were permitted to call time-outs or terminate the experiment at the onset of discomfort. Prior to beginning the experiment, subjects completed the Hidden Figures Test, Immersive Tendencies Questionnaire, and Simulator Sickness Questionnaire. Once they were comfortably situated in the helmet, the Postural Stability test was conducted for 30 seconds before the experiment began.

The postural stability test was conducted with the helmet worn and the visuals turned off, providing an automatic visual blackout condition in the already darkened room, and allowing head tracking during the task. Subjects were required to cross their arms over their chest and place their hands on their shoulders. They were then required to stand on one foot for thirty (30) seconds, with the opposing foot simply lifted off the floor and not touching the standing leg. They were instructed that if they needed to they could stabilize themselves by returning the lifted foot to the floor for a second. These foot touches were recorded. A spotter was always provided for safety during this measure.

All participants remained seated during the performance of the tasks. **Bins, Fixed Tracking,** and **Moving Tracking** task sequences required the subject to repeat the task ten times. The **Distance Estimation** task required responses at five different distances during the task sequence. **Doorways** required crossing ten rooms during the task sequence. Each task sequence was performed three times, with an enforced five minute break between sets of tasks (a set being one sequence of each task). **Distance Estimation** was always performed first, with the other tasks presented in counter-balanced order such that each task followed each of the others an equal number of times across subjects.

During the performance of the tasks, subjects were generally and unconditionally reinforced by the experimenter (e.g. "good job" or "good effort"). Following the completion of the final set of tasks, the Postural Stability test was again performed. Finally, the Presence Questionnaire (PQ), another Simulator Sickness, and a VR effectiveness questionnaires were administered at the end of the session. All subjects remained for a minimum of approximately 30 minutes after the completion of the third set of tasks, during which the questionnaires were completed, to assure no short term aftereffects of exposure to the Virtual Environment.

Results

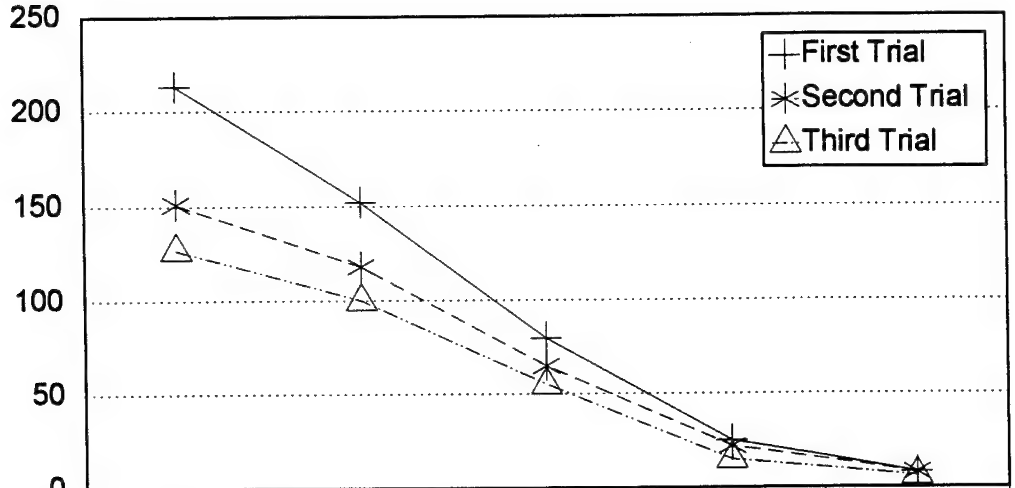
In general, the means of all dependent measures for the performance tasks were calculated over 5 trial segments, resulting in six repeated measures (means) for each task. This

was done in order to alleviate some of the trial-to-trial fluctuation resulting from the random placement of the target in the tasks and to prevent too many "missing value" cases from early in the trials. Repeated measures ANOVA's were performed on these segment means for each experimental task to determine the effects of the major independent between-subjects variables VIEW (stereoscopic or monoscopic) and COUPLING (head-tracked or non-head-tracked).

Distance Estimation

A repeated measures MANOVA was conducted on two derivatives of the actual distance estimates made by the subjects (five distances estimated during three trial sessions, see the **Distance Estimation** task description, above); the Error at each distance (subject estimated target distance - actual object distance at judgment) and the Percent Error ([subject estimated target - actual distance] / target distance). For example a subject might judge that the approaching figure was thirty (30) feet away when it was actually thirty-three (33) feet away, leading to an error of negative three or a percent error of negative ten percent. Thus negative errors indicate that subjects underestimated the actual distance (which was further away than the to-be-estimated distance). Negative percentages indicate the same thing, and provide the additional scale information of how much underestimation there actually was (negative 100% indicating that the object was twice as far away as the to-be-estimated distance). For both of these response measures, there was a significant main effect for COUPLING and DISTANCE. The analysis showed that Head-Coupling was better than no Head-Coupling (Error: $F=5.19$, $p<.028$, Coupled $M=-3.99$, Uncoupled $M=-5.32$; Percent Error: $F=4.36$, $p<.043$, Coupled $M=-65.36\%$, Uncoupled $M=-87.56\%$).

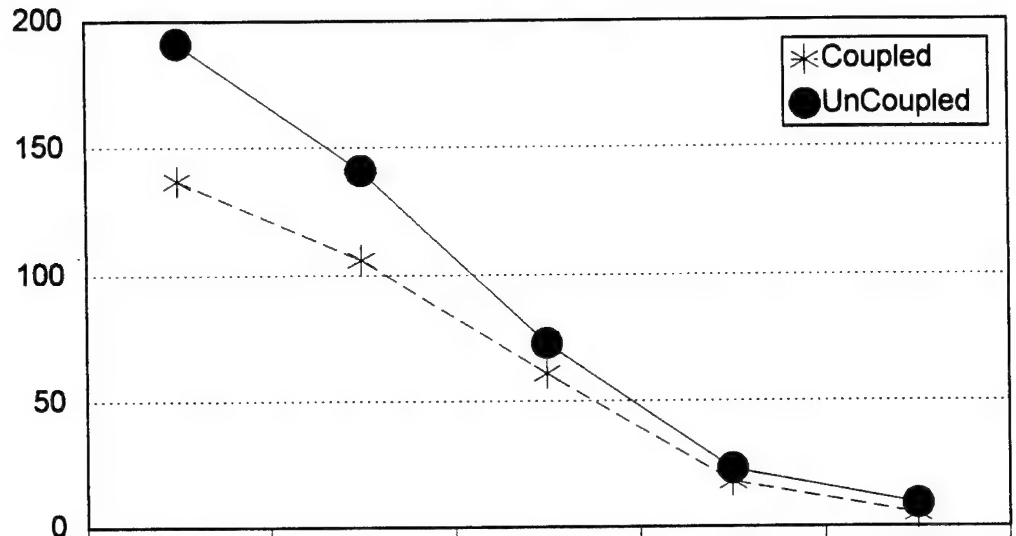
The analysis of the within-subject DISTANCE factor found significant differences for estimation of DISTANCE (Error: $F=79.02$, $p<.001$; Percent Error: $F=109.68$, $p<.001$). Subjects seemed to be better at estimating distances farther away, as indicated by the changes in means (M@2.5=-163.7%; M@5=-123.7%; M@10=-66.7%; M@20=-20.7%; M@30=-7.3%). Post Hoc analyses on these distances revealed that there were significant differences in the accuracy of estimates between: thirty and ten, five, or two-and-one-half feet; twenty and five or two-and-one-half feet; and ten and five or two-and-one-half feet ($S=50.4$, $p<.05$). The error means are mathematically related and present the same pattern.



	2.5	5	10	20	30
First Trial	213	152	80	25	8
Second Trial	151	118	65	22	8
Third Trial	127	100	55	15	6

Figure 1. Mean percent error for distance estimates by trial.

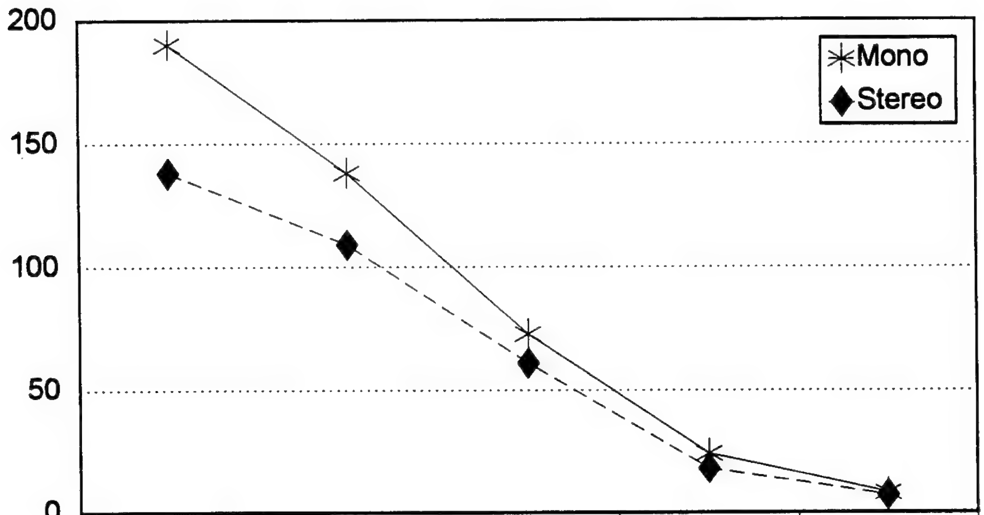
The MANOVA also revealed that subjects improved significantly over trials at judging distances (F=16.89, p<.001), and improved differentially, as shown by the significant interaction found between DISTANCE and TRIAL with the Percent Error measure (F=31.44, p<.001). This interaction is shown in the means presented in Figure 1, but was not clarified by any straight-forward post hoc contrast (using Scheffe's statistic). The non-statistically supported interpretation of the means indicates that people were improving at judging short distances over the three trial sessions.



	2.5	5	10	20	30
Coupled	137	106	61	18	5
UnCoupled	191	141	73	23	9

Figure 2. Mean percent error for distance estimations by coupling condition.

A significant main factor interaction between COUPLING and VIEW was also found with the Error measurement (F=5.13, p<.028, Stereo/Coupled =-6.71, Mono/Coupled =-12.9, Stereo/Uncoupled =-13.2, Mono/Uncoupled =-13.07). This main effect interaction is averaged across the distances judged, and is therefore less informative than significant interactions that incorporate the range of distance estimations would be. Still, the analysis indicates that combining stereoscopic presentations with head-coupling significantly improves the judgment of relatively short distances of thirty feet or less.



	2.5	5	10	20	30
Mono	190	138	73	24	8
Stereo	138	109	61	18	7

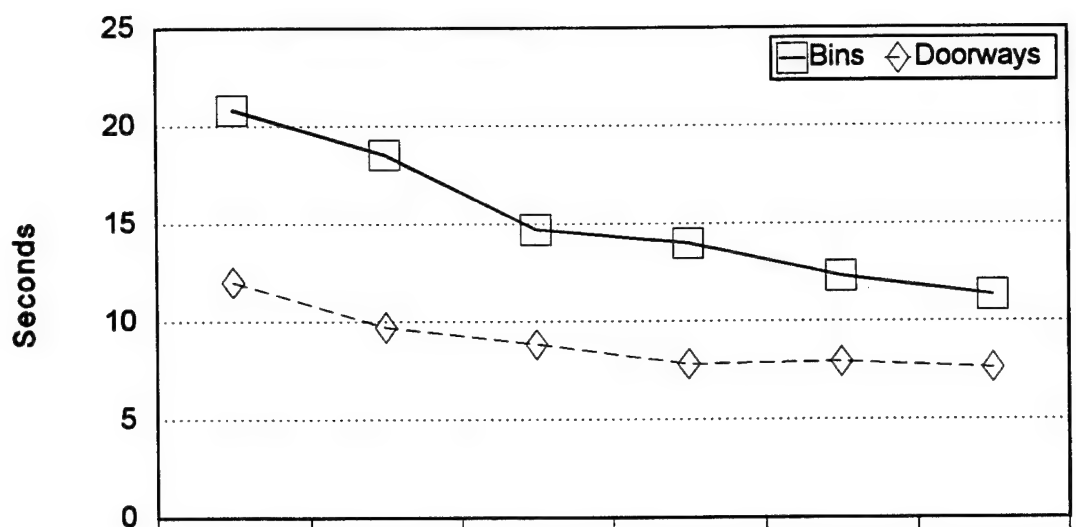
Figure 3. Mean percent error for distance estimates by viewing condition.

The repeated measures MANOVA using <u>Percent Error</u> also revealed significant interactions between DISTANCE and COUPLING ($F=2.93$, $p<.022$, see Figure 2 for means), and between DISTANCE and VIEW ($F=2.70$, $p<.032$, see Figure 3 for means). This seems to indicate that Head-Coupling allowed more accurate estimation at shorter distances than Non-Head-Coupling (see Figure 2), and Stereoscopic viewing also was improved at shorter distances over Monoscopic presentations (see Figure 3). However, post hoc analyses did not reveal any significant contrasts (relative to factor and group means, simple contrasts, or linear contrasts) where this significant interaction was occurring (Marascuilo & Levin, 1970; Kirk, 1968). This leaves us without a clear statistical interpretation of the interaction between these factors. However, a non-statistically supported visual interpretation of the means presented in Figures 2 and 3 indicates that both Stereoscopic presentations and Head-Coupling as separate factors contributed to a decreased amount of error in distance estimation at the relatively short distances used in this Virtual Environment.

Bins

The repeated measures MANOVA conducted for performance Time in the **Bins** task found significance for the repeated TRIALS only ($F=38.76$, $p<.001$). The means for Bins performance over the trial segments are presented with the **Doorways** task means in Figure 4. The **Bins** and **Doorways** data are presented together because they have similar patterns of results and are both measured in terms of time to perform the task successfully.

A post hoc analysis on the TRIALS for **Bins** found significant differences in performance times between the first segment and the third, fourth, fifth, and sixth segments (Scheffe, $S=4.039$, $p<.05$). There were also significant differences between the second segment and the fourth, fifth, and sixth segments (see Figure 4). This indicates that considerable learning was occurring during the initial trials, and that some performance ceiling was perhaps being reached during the later trials.

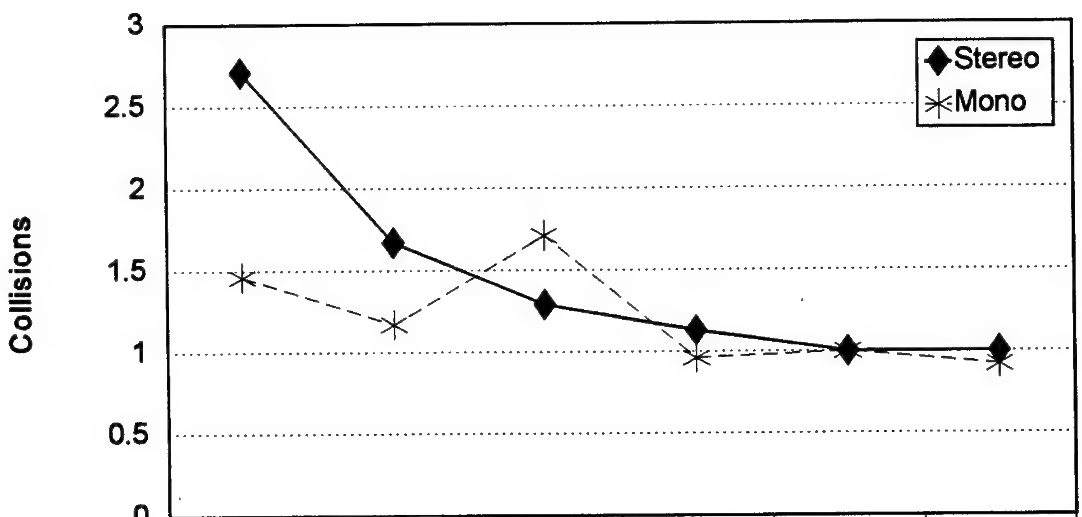


	1-5	6-10	11-15	16-20	21-25	26-30
Bins	20.81	18.52	14.69	13.97	12.3	11.33
Doorways	12.01	9.71	8.83	7.83	7.93	7.63

Figure 4. Performance times for bins and doorways across repeated trials.

Doorways

Performance in the **Doorways** task was measured by the Time to cross and exit each room (using the mean for five consecutive rooms) and by the number of Collisions made (again, the mean for five rooms, see Figure 4). The significant results of the MANOVAs on both measures showed that subjects improved in their performance over repeated TRIALS (Time: $F=331.57$, $p<.001$; Collisions: $F=7.44$, $p<.001$). The repeated TRIALS were not homogeneous, but the Greenhouse-Geisser adjustment of degrees of freedom (to 2.84 from 5) also found the repeated TRIALS significant (Kirk, 1968) which indicates that the results should be accepted. The mean performance times for **Doorways** over the trial segments are presented in Figure 4, in combination with the **Bins** data.



	1-5	6-10	11-15	16-20	21-25	26-30
Stereo	2.71	1.67	1.29	1.13	1	1
Mono	1.46	1.17	1.71	0.96	1	0.92

Figure 5. Mean number of collisions per doorways segment by viewing condition.

The number of Collisions was also non-homogeneous across trials, being positively skewed with a large number of zeros entered as data. A logarithmic transformation was applied after

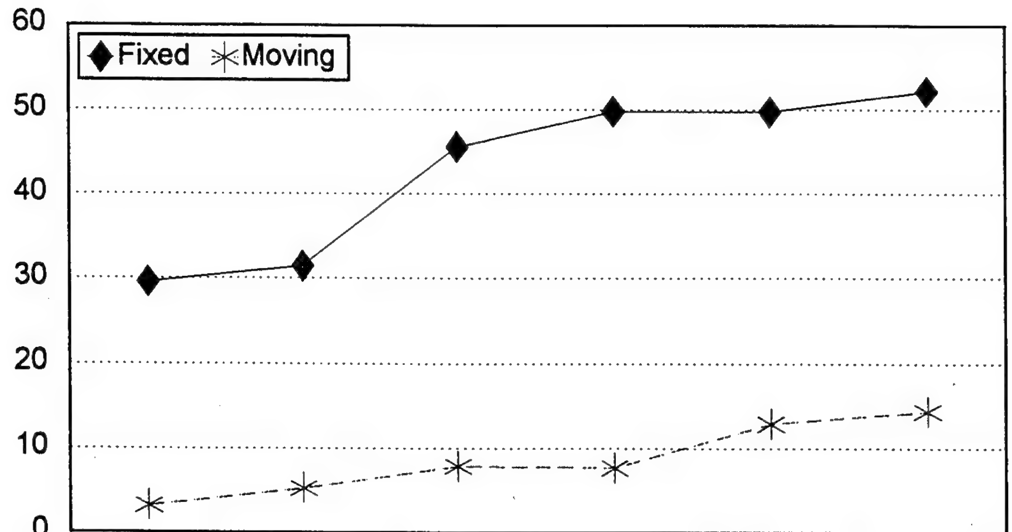
adding one to each cell value (a method suggested by Kirk, 1968 to correct for data of this type) to normalize the data, and a MANOVA was conducted. The repeated TRIALS remained significant ($F=5.22$, $p<.001$) and an interaction was also revealed between VIEW and TRIALS ($F=4.44$, $p<.001$). The interaction is depicted in Figure 5 using the original collision means. As can be seen in the figure, the stereoscopic condition produced many more collisions initially, although with practice the performance became very similar in each condition.

Post Hoc analyses were performed on contrasts between the means (normalized as described above) based on the apparent interactions in the unadjusted means. The difference between the mean Collisions in the first set (rooms 1-5) versus the third set (rooms 11-15) for the Stereoscopic view was compared with the same contrast (set 1 - set 3) for the Monoscopic View. This comparison was significant (Normalized Mean Difference =.3542, $S=.2746$, $p<.05$) as might be expected from an inspection of Figure 5 (showing unadjusted means). The difference between mean Collisions with Stereoscopic and Monoscopic presentations in the first (five room) segment also contrasted significantly with the difference between mean Collisions with Stereoscopic and Monoscopic presentation in the third set of rooms (Normalized Mean Difference =.35415, $S=.2746$, $p<.05$). None of the other contrasts were significant.

Fixed Tracking

The measures in the tracking tasks are different, as noted in the task descriptions above and shown in Figure 6. The maximum time allowed to perform each trial was fixed, with three measures being collected and analyzed. The three measures used in analyses were Percent Time on Target, Time to Target, and Misses (unsuccessful trials). The means for Time to Target do not reflect the unsuccessful trials, but the Percent Time on Target means do reflect the unsuccessful trials. The only significant finding for **Fixed Tracking** with these measures was the repeated TRIALS (Percent Time: $F=45.73$, $p<.001$; Time to Target: $F=15.46$, $p<.001$; Misses: $F=23.81$, $p<.001$). Although these repeated measures were non-homogeneous, adjustments to the degrees of freedom (the Greenhouse-Geisser adjustment; see Kirk, 1968) supported these results as significant.

The finding that people get better with practice is not surprising. The Percent Time on Target trial means for **Fixed Tracking** presented in Figure 6 provide a clear picture of the significant results. A post hoc analysis of the TRIALS (using Percent Time on Target) found significant differences between the first segment and the third, fourth, fifth, and sixth segments (Scheffe $S=10.79$, $p<.05$). The analysis also found significant differences between the mean for the second segment and the third, fourth, fifth, and sixth segment means.



	1-5	6-10	11-15	16-20	21-25	26-30
Fixed	29.59	31.47	45.57	49.79	49.75	52.2
Moving	3.21	5.25	7.87	7.7	12.92	14.39

Figure 6. Mean percent time on target over repeated trials for FIXED and MOVING tracking tasks.

<u>Moving Tracking</u>

The **Moving Tracking** task was measured with the same three types of scores as **Fixed Tracking** in repeated measures MANOVAs. As with **Fixed Tracking**, the repeated TRIALS were also significant for **Moving Tracking** (<u>Percent Time on Target</u>: $F=25.04$, $p<.001$; <u>Misses</u>: $F=20.71$, $p<.001$), as might be expected. The <u>Time to Target</u> results were unusable due to missing data, which left the cells of the analysis unbalanced. The means for **Moving Tracking** over the TRIAL segments are also presented in Figure 6.

A post hoc analysis (Scheffe, Kirk, 1968) using the <u>Percent Time on Target</u> data revealed that significant changes in performance were occurring over the segments or trials. Significant differences were found between the mean performance scores on the first segment when contrasted with the fifth and sixth segments, and the second segment when contrasted with the fifth and sixth segments ($S=5.91$, $p<.05$). The sixth segment mean

was also significantly different from the third and fourth segment means. The seemingly delayed acquisition pattern shown in these results may indicate that no ceiling on performance had been reached for this task.

One significant interaction with the repeated measures was discovered in the analysis of the **Moving Tracking** task. The VIEW X TRIALS for Percent Time on Target was statistically significant, $F= 2.28$, $p < .048$. However, several of the repeated measures cells were significantly non-normal, and therefore the Greenhouse-Geisser degrees of freedom adjustment was checked (as recommended by Kirk, 1968) and did not indicate a significant interaction. As Kirk points out, whenever the F statistic is significant and the Greenhouse-Geisser is not, Hotellings T^2 should be checked as the better measure of significance (Kirk, p. 262). In this case, the Hotellings statistic did not indicate significance, and no further analyses were conducted.

Simulator Sickness

A repeated measures MANOVA (pre-experiment vs. post-experiment) was performed on the Simulator Sickness Questionnaire (SSQ) total and subscales. Values for subscales on Nausea, Oculomotor, and Disorientation, as well as Total Severity were calculated using Kennedy, Lane, et al (1993) scales and methods. Analyses revealed overall significant increases in Total Severity ($F=8.59$, $p<.005$), Nausea ($F=4.43$, $p<.041$), Disorientation ($F=6.41$, $p<.015$), and Oculomotor ($F=8.17$, $p<.006$). The means for all the scales are presented in Table 1.

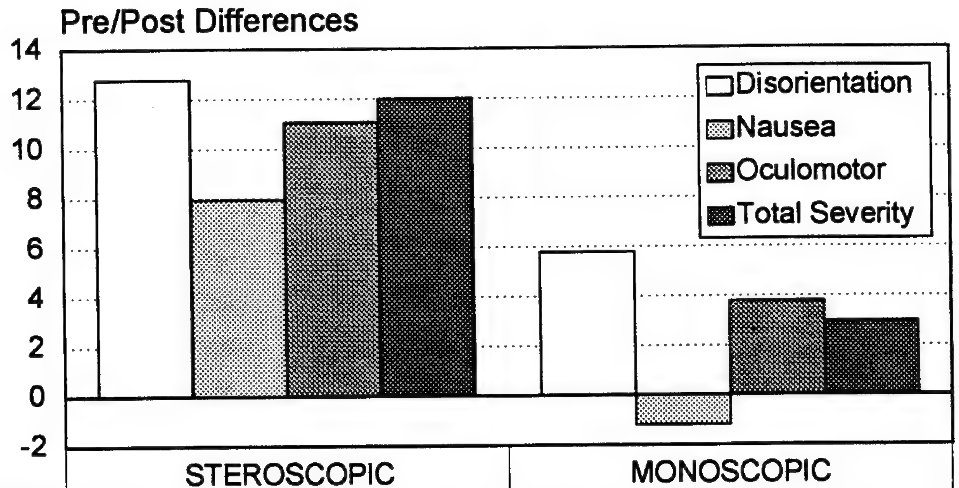
Table 1

Pre- and Post-Experiment Simulator Sickness Means

	Pre	Post	Difference
Total Severity	10.05	17.53	7.48
Nausea	7.15	10.53	3.38
Oculomotor	11.37	18.79	7.42
Disorientation	6.09	15.37	9.26

In addition, there was a significant interaction between the VIEW conditions within the pre- and post-experimental Nausea subscale values ($F=8.12$, $p<.007$). Although both groups (Stereoscopic and Monoscopic) started with a measured Nausea level of 7.15, the mean level after Stereoscopic presentation was 15.1 while the mean after Monoscopic presentation was only 5.96 (a decrease in symptoms, see Figure 7). No other significant

interactions in the pre- and post-experimental administrations of the SSQ were revealed.



	STEROSCOPIC	MONOSCOPIC
Disorientation	12.76	5.8
Nausea	7.95	-1.19
Oculomotor	11.05	3.79
Total Severity	12	2.96

Figure 7. Mean Differences between pre- and post-experiment Simulator Sickness Questionnaires by viewing condition.

Other analyses (Correlations and MANCOVAS) explored the possibility of either pre- or post-experimental SSQ subscale scores relating to the performance of the experimental tasks without any notable success. There were a few significant correlations between performance in the different segments of the tasks and the SSQ measures. Given eight measures (four pre-and post) and six performance segments, there are forty-eight correlations to consider for each task. While there were significant correlations with some of the subscale scores and even a few for the task performance scores on some trials, only one significant relationship was suggested between subscales and repeated task performance. The post-experimental Disorientation and Oculomotor subscales and the Total Severity measure all correlated significantly (positive, between $r=.29$, $p<.044$ and $r=.39$, $p<.006$) with the second, fourth, and sixth segments of the **BINS** task. These segments are the final trials in the repeated sequences (see task description, above). The investigation of

these scales as covariates in analyzing the performance of the task resulted in no changes to previous findings.

Postural Stability

Postural stability was measured using the head-tracking devices to measure the position of the head in six degrees of freedom approximately every 0.54 seconds. These data were then used to generate the mean deviation of head movement from the initial starting point during the thirty (30) second task. The standard deviation of distance from the starting head point was also calculated. These measures were intended to provide an indication of the amount of sway experienced by each individual. Tests of homogeneity revealed that the cells were not normal, so a log transform was applied to the data. A repeated measures ANOVA was performed on the transformed pre- and post-experiment postural stability measures (both mean distance and the standard deviation of the distance). No within subjects effects were revealed for these measures, indicating that there was no overall change between the pre- and post-experiment sessions. There were no significant effects associated with the Between-Subjects experimental variables.

Presence

The Immersive Tendencies Questionnaire(ITQ) and Presence Questionnaire (PQ) totals and subscales were generated based on the latest structuring for these developing scales (Singer, Witmer, & Bailey, 1994). The data gathered in this experiment were included in that analysis, which selected for items that would provide optimal reliability values for the scale totals and cluster analyses for the subscales. The reliability (Cronbach's alpha) for the scales found in that analysis was .81 for the ITQ Total and .88 for the PQ Total. The correlation between the ITQ Total and PQ Total was .347 (Pearsons r, p<.016). The ITQ subscales (Games: involvement in video games; Involvement: the tendency to ignore external stimuli in some situations; Focus: current mental alertness and concentration factors) all correlated significantly with ITQ Total, and not with each other, as expected from the cluster analysis. The PQ is more of a unitary scale, and all subscales (Involved/Control: involvement and VE control/responsiveness issues; Natural: VE reality and naturalness; Interface Quality: VE control/display interference in the activity) correlated significantly with one another as well as with the PQ Total. The methods used in the determination of subscale structures are not germane to this experiment and will not be further presented or discussed. The methods are thoroughly discussed in Witmer and Singer (1994) or Singer, Witmer, and Bailey (1994) in conjunction with a more detailed discussion of the Presence phenomenon.

Table 2

Immersive Tendencies & Presence Questionnaire means: Totals and subscales by experimental condition

	TOTAL	ST/CPLD	ST/UNCPLD	MONO/CPLD	MONO/UNCPLD
ITQ TOTAL	76.56	75.75	80.42	75.18	73.67
INVOLVEMENT	25.60	25.33	25.75	26.42	24.92
FOCUS	35.06	34.83	36.58	33.92	34.92
GAMES	6.33	6.0	7.33	7.58	4.42
PQ TOTAL	92.58	88.25	99.00	99.36	82.75
INVLVD/CNTRL	58.69	56.17	60.67	63.45	53.83
NATURAL	11.94	10.83	13.08	12.91	10.42
INTERFACE	14.42	13.00	16.67	15.64	12.75

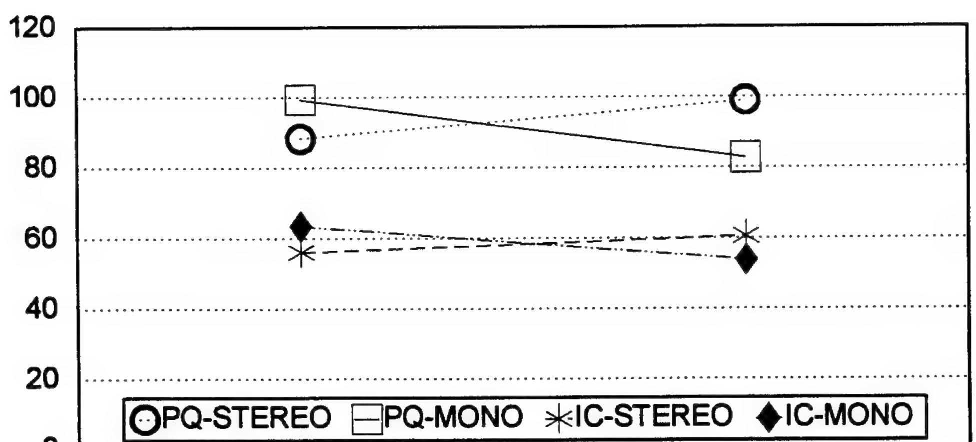
ITQ and Experimental Factors. A MANOVA was performed using the ITQ scales, to check for possible immersive tendencies that could have biased the experiment. The Games subscale analysis revealed a COUPLING X VIEW interaction ($F=7.80$, $p<.008$; see Table 2 for Means). This indicated that there was something about the subjects previous experience or behavioral tendencies that was different for the different conditions. The Games subscale addresses past experience with video and arcade games, so the interaction indicated a possible bias that could have differentially influenced the performance of the experimental tasks in the different conditions. As indicated by the means in Table 2, subjects with the highest scores were in the STEREO/UNCOUPLED and MONO/COUPLED conditions.

At this point in the experimental sequence, statistical control is the only method available to investigate this potential confound (Kirk, 1968). Correlations were conducted with the performance measures for all the tasks, with very few significant correlations appearing. (The comparisons with performance measures for the tasks were treated as a family for analyses, with alpha set to .007 for a resultant family-wide error rate of .048.) A series of MANCOVAs were also run for each of the experimental tasks using both Games and ITQ Total (as GAMES correlates significantly with ITQ Total). These analyses resulted in no changes to the significance effects for between subjects factors of COUPLING, VIEW, or their interaction for **Bins**, **Doorways**, **Fixed Tracking**, or **Moving Tracking**. These analyses show that the different distribution of scores on the Games subscale had no effect on the experiment results.

ITQ and Simulator Sickness. There are sixteen correlations possible between the ITQ and the pre-experiment administration of

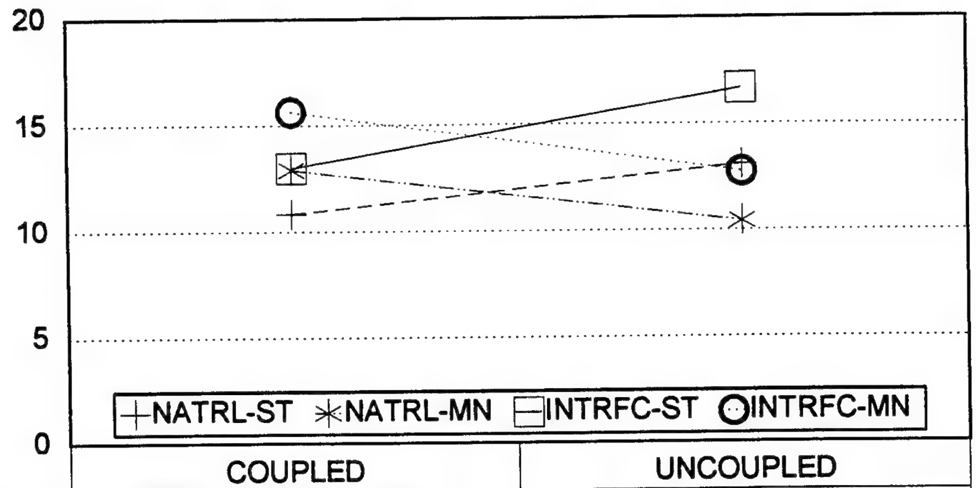
the SSQ, and another sixteen possible between the ITQ and the post-experiment administration of the SSQ. One half of the correlations between the ITQ and the pre-experiment SSQ were negative. However, there was only one significant correlation in the pre-test set, between the <u>Games</u> subscale and the <u>Oculomotor</u> subscale (r=.3125, p<.032). Ten out of the possible sixteen correlations between the ITQ scales and the post-experiment SSQ administration were negative. Again, there were only two significant relationships found between the ITQ scales and the post-experiment SSQ scales. The ITQ scale <u>Focus</u> was significantly negatively related to both <u>Oculomotor</u> (r=-.37, p<.011) and <u>Total Severity</u> (r=-.33, p<.023). These results indicate little in terms of a consistent relationship between these measures.

<u>ITQ and Hidden Figures</u>. A correlation between the <u>Hidden Figures</u> test and the ITQ scales found no significant relationships. Further analyses were not pursued.



	COUPLED	UNCOUPLED
PQ-STEREO	88.25	99
PQ-MONO	99.36	82.75
IC-STEREO	56.17	60.67
IC-MONO	63.45	53.83

Figure 8. Means for the Presence Questionnaire and the Involved/Control subscale by experimental condition.



	COUPLED	UNCOUPLED
NATRL-ST	10.83	13.08
NATRL-MN	12.91	10.42
INTRFC-ST	13	16.67
INTRFC-MN	15.64	12.75

Figure 9. Means for Naturalness and Interface Quality subscales by experimental condition.

PQ and Experimental Factors. A MANOVA was conducted on the PQ Total and subscales based on the experimental conditions. There was a significant VIEW x COUPLING interaction ($F=12.22$, $p<.001$) for the PQ Total (see Figure 8). All of the subscales also had significant interactions (Involved/Control: $F=11.23$, $p<.002$; Natural: $F=6.71$, $p<.013$; Interface Quality: $F=10.02$, $p<.003$; see Figures 8 & 9). These results show that different levels of VR factors are experienced differently by subjects.

PQ and Performance. Correlations were performed using the PQ scales and the performance outcomes of the experimental tasks to investigate the relationship between presence and task performance. Given the number of correlations possible between the performance measures (repeated) for any individual task and the four measures of presence, an adjusted significance level was used as the criteria (with the family of comparisons being the performance measures times the Presence scales, 28 comparisons [seven per task X four tasks] each at $p<.001$ provides an overall family of comparisons significance level of .0276; Kirk, 1968).

The PQ Total and Subscales did not correlate significantly with the performance measures for the four tasks.

PQ and Simulator Sickness. There were no significant correlations between the PQ total and subscale scores and the post-experience Simulator Sickness Questionnaire total and subscale scores (again using alpha=.003 for each comparison to get alpha=.0469 for the family of comparisons; Kirk, 1968). Fifteen out of sixteen possible correlations between the scales (four SSQ scales X four PQ scales) were negative, as has repeatedly been the case in other research (Witmer & Singer, 1994; Singer, Witmer, & Bailey, 1994). This finding also held true for the correlation between the SSQ differences (Pre - Post) and the Presence measures, with fourteen of sixteen being negative correlations (using the same alpha level adjustments).

Table 3

Mean Hidden Figures scores for Field Dependence by experimental condition.

	FIELD	
	DEPENDENT	INDEPENDENT
STEREO/COUPLED	8.33 (9)	18.67 (3)
STEREO/UNCOUPLED	7.88 (8)	20.25 (4)
MONO/COUPLED	8.67 (9)	18.67 (3)
MONO/UNCOUPLED	9.00 (9)	21.67 (3)

Number per cell in parentheses

Cognitive Style

The Hidden Figures test was used to investigate the "types" of cognitive processes evident in the subjects assigned to the different cells. Subjects were assigned to the different conditions based on their scoring on the test as Field Dependent or Field Independent. The means by condition are presented in Table 3. As can be seen in that table, the number of Field Independent subjects was low and a roughly proportional assignment scheme was used. The scores on the test did not correlate with the totals or subscales of the Immersive Tendencies Questionnaire, the Presence Questionnaire, or the Simulator Sickness Questionnaire. The Hidden Figures scores did correlate significantly with the performance scores for the **Bins** task, but a MANCOVA did not reveal any significant changes to those results.

Discussion

The objective of this experiment was to investigate the relative effectiveness of stereoscopic and monoscopic displays, in combination with the effects of head-coupled or non-head-coupled display control. Results indicate no significant main or interactive effects for VIEW or COUPLING conditions in **Bins**, **Doorways**, **Fixed Tracking**, or **Moving Tracking** tasks. The most interesting and informative results were found in the significant main and interaction effects revealed in the **Distance Estimation** task.

Distance Estimation

Because no overt or informative feedback was given to the subjects when estimating distances other than the first distance (forty feet), the improvement in distance estimation over trials indicates some kind of adaptation and/or spontaneous learning. There was implicit informative feedback for both the longest distance and the shortest distances. After the first trial, most subjects knew that the object at the end of the hallway was a six foot tall male standing forty feet away (and after the second trial, everyone caught on). Given the familiarity of the target, it may become easier to "hit" the first distance (thirty feet), knowing the start point of forty feet and using a movement timing strategy. At the other end, by the second and third repetitions the short distances may have become easier because that is where the soldier's features become more detectable, which could combine with movement timing strategies to provide more accurate judgments of distance. Given one's knowledge of what the soldier target looked like, it may also have become easier to find visual cues for the ten, five, and two-and-a-half foot distances. In addition, learning (over the three trials) the time required for traversing the entire forty feet may have provided additional information for the incorporation of timing estimates.

There are still large errors made at shorter distances, however. The discrepancy may have been influenced by the field of view provided by the HMD (83° degrees horizontal and 41° degrees vertical). The limited field of view could distort some of the normal cues used to estimate short distances, especially since the judgements were required while the target was moving toward the subject from a known distance of forty feet. The effect is like cupping your hands around your eyes, which causes a large object (like our target soldier representation at close distances) to look larger by filling the available field of view. As an example, at two-and-a-half feet with a restricted field of view, only the head and shoulders of a human figure can be seen and the figure fills the visual field. Some of the subjects would actually flinch when the figure came within this distance, as if they were about to collide (the image would pass through

the viewpoint at the end of the trial, usually a half second later).

The significant interaction between the distances estimated and COUPLING, and between distances estimated and VIEW is interesting and informative. First, because the improvement is with Stereoscopic over Monoscopic displays and with Head-coupling over Non-head-coupling, which helps to explain the significant overall interaction between VIEW and COUPLING found in this data. Second, the improvement in accuracy with these conditions follows our expectations, that a more natural interaction would support improvements in performance. The Stereoscopic presentation provided some cues at shorter distances that improved the estimations. Presumably these are the larger disparity and convergence cues that were eliminated from the monoscopic presentation. The improvement with Head-coupling is even easier to explain. Even with monoscopic presentations, at short distances the use of coupling allows small head movements to provide parallax cues for the approaching figure. The pattern of inaccuracies at these distances indicates problems for veridical perception in VEs. The large errors still occurring at short distances indicates that VE presentation technology requires considerable improvement.

Movement and Manipulation Tasks

The lack of consistent and large scale significant results for VIEW (Stereoscopic / Monoscopic presentation) or for COUPLING (head- or non-head-coupled) with the movement or manipulation tasks is unexpected. Based on the literature, we expected the stereoscopic display to result in some consistent performance gains over monoscopic displays. There are several possible explanatory factors for this finding. The first factor may be that the tasks were performed over distances that decreased the importance of the stereoscopic information that could be presented in the HMD.

While pilot testing the experiment, it was discovered that on the **Fixed Tracking** task, the target frequently appeared out of view in the non-head-coupled condition. In the non-head-coupled condition, the limited field of view produced a truncated area in which the target could appear. Similarly, in the **Moving Tracking** task, the target might not enter the view for a relatively long time or might not be visible at the end of the trial, which would create a significant difference between the head-coupled and non-coupled conditions. (The software generated the target at a wall and then moved it to an opposing surface. In both tasks, the location or path of the target was randomly generated.) Therefore, to make the conditions as comparable as possible, the room was lengthened from ten to twenty feet. This placed the viewer 19.5 feet from the most distant plane on which the target would appear. We then imposed a limit on how close the target

could begin, so that the target never started closer than six feet from the subject's point of view. Even with this change, thirty trials (out of 1440, 2.08%) in the fixed, and ten trials (.69%) in the moving tracking still had to be repeated during the experiment, because the target ball never entered the field of view. We believe, in retrospect, that lengthening the room served to increase the availability of two dimensional cues (such as texture gradient information) while lessening the salience of the stereoscopic cues. As a result of the extra distance, the tracking tasks may have been pushed into a range in which the importance and influence of stereoscopic cues became minimal. In addition, the room change also served to decrease the information available from head movement.

This point also holds for the **Doorways** task, although not as well for the **Bins** task. In **Doorways** the subject was crossing rooms that were approximately twenty feet square, with the major goal being to pass through a doorway that was represented as normal sized (approximately three feet wide and seven feet tall). As the door locations were staggered, the longest viewing distances were seldom more than twenty-eight feet. Obviously, the final alignment for passing through the door would occur at a relatively short distance. The significant interaction between the presentation mode and repeated trials for collisions presents an interesting pattern. The Stereoscopic presentation is initially high and shows a classic adaptation curve while the Monoscopic presentation mode starts lower and shows more problems in the middle segments before matching the Stereoscopic performance levels in the last two trials.

Our results seem to indicate that the extra computing time and computer memory requirements for generating stereoscopic displays may not be necessary for tasks performed at "middle" distances (from nine to sixty feet). This is, unfortunately, an inference supported by the "not significantly different" results we have found in this experiment. As all researchers know, and the consumers of research should know, not significantly different does **not** mean there is no difference. It remains to be seen whether or not very close work (within several feet) would show an effect for stereoscopic vision in virtual environments. These distances are the very ones at which current display technology performs most poorly.

Presentation Differences

In some experiments testing monoscopic vision (Arthur, Ware, & Booth, 1993), one eye is occluded. However, when one eye is occluded, the scene shifts in the direction of the non-occluded eye (Ono, 1991). Consequently, the ball that one interprets as being directly in front, has now been shifted left or right. When the individual reaches or points, the object simply is "not in the right place." Therefore, performance decrements for

monoscopic displays in some previous studies may be due to the shifting of the scene.

The results of our experiment may be an outcome of the way in which the monoscopic condition was rendered. In our monoscopic condition, parallax was set to zero, so that each eye received the same picture. The convergence angle was subsequently changed, so that the views overlapped (an adjustment that was supposed to match the interpupillary distance of the subject). In comparison, Sollenberger and Milgram (1993) used a flat screen as a monoscopic condition, while the stereoscopic conditions were produced via special glasses which alternately occluded the left and right eye, at a rate fast enough to preclude conscious detection. In addition, these glasses can overemphasize the Inter-Ocular Distance and over-exaggerate normal stereopsis. Consequently, their conditions for stereoscopic and monoscopic viewing were not exactly equivalent, and the results may be from the different set-up used for each condition. Research is currently being conducted at the U.S. Army Research Institute, Simulator Systems Research Unit to determine if there are any performance differences in flat screens vs. HMDs (see Lampton, Knerr, Goldberg, Bliss, Moshell, & Blau, 1994).

Finally, several subjects commented on the poor quality of the stereoscopic effect. These subjects compared these displays to the relatively poor three dimensional effect generated by 3D movies. In such films, binocular disparity is exaggerated far beyond normal disparity (Dr. Milton Katz, personal communication, 1994), and this was not the case in our equipment configuration.

Simulator Sickness

Perhaps the most interesting findings of this experiment were the simulator sickness results. A significant increase in sickness after exposure to the VR was not expected, as we had tried to alleviate the problem in order to maintain experimental efficiency. However, based on findings by Bailey (1994), we were prepared for a significant effect for Head-coupling in this regard. It was most interesting to find a significant within-subjects effect for the VIEW condition. Specifically, the stereoscopic condition was more nauseogenic than the monoscopic display. The reasons for this remain unclear. The presentation system (see above) was set up to compute views for both eyes in both stereoscopic and monoscopic conditions. Therefore, no difference in lag time, which has previously been found to be nauseogenic, existed between the two conditions used in this research.

The modifications used to get stereoscopic presentations may be the cause of the nauseogenic interaction. As discussed in the introduction, stereoscopic vision is based on the disparate views

seen by each of our eyes. In the stereo HMD, the computer generates a scene from the data and projects two separate but overlapping views (one to each lens in the helmet). These views are defined based on the parallax values (essentially the Inter-Ocular Distance, or distance between the pupils [also referred to as the Inter Pupillary Distance]) and aligned by the convergence adjustment performed for each subject using the generating software. When the stereoscopic view was displayed in front of each eye, the disjuncts discussed in the introduction occur. The monoscopic presentation set the parallax to zero, projecting the <u>same</u> view to each eye. However, nobody's IOD was set accurately in those conditions, yet the monoscopic conditions turned out to be less nauseogenic than the stereoscopic. In both conditions, the convergence then aligned the views for the subjects' preferred focal distance. The alignment in the monoscopic presentation may have served to decrease any remaining disjunct, which was still experienced between accommodation and vergence in the stereoscopic presentation. Alternatively, the problem could be that the average parallax used in the stereoscopic condition, when adjusted for each individual's convergence, exacerbated the problem for subjects in that condition. It may be that simplifying the complexity of the display, in the monoscopic condition, eased the interpretational load on the subjects, resulting in less nausea.

There is one other simulator sickness issue that arose towards the end of this study; the phenomenon of delayed simulator sickness. A few subjects reported experiencing minor headaches, tiredness, or flashbacks several hours after their exposure to the virtual environment. Unfortunately, by the time we had heard of such problems, several months had passed since the earliest subjects had participated. Many had moved away at the end of the semester, so we were unable to contact them. Of those we contacted, most did not report any symptoms. Future research should continue in this area, with scheduled follow-up interviews the day after exposure to the virtual environment. In general this is a known phenomenon, but little researched and poorly understood issue (Murphy, 1993; Brown & Baloh, 1987; Gordon, Spitzer, Shupak, & Doweck, 1992).

Presence and Immersive Tendencies

The initial finding with the GAMES subscale from the Immersive Tendencies Questionnaire was somewhat disconcerting. The GAMES subscale, as mentioned above, taps experience with video and arcade games. This could be a directly confounding area for many investigations where controls or displays are similar to arcade or video games. The analyses showing that there was no significant change to the outcome of the major variables based on the subscale confirmed the balanced assignment of subjects to conditions.

The lack of relationship between the subscale and performance under the conditions shows that there was no confounding in this instance. This finding is both good and bad. If a relationship had been found between game experience and performance of the tasks, an adjustment would have been necessary to correctly analyze the effects of the experimental conditions. Since no relationship was found with task performance, no adjustments were necessary. On the other hand, a relationship between gaming experience and initial performance or early learning patterns would have provided information that could have been used to predict entry level skills for these types of tasks in VR situations. A more thorough development of this subscale, possibly with additional questions, may be of value in order to balance assignment to VR conditions in future research, and if relationships can be found, to adapt VR-based training to take advantage or prior skills and knowledges.

Cognitive Style

The Hidden Figures test (Ekstrom, et al., 1976) measure of Field Dependence did not relate in any significant way with any of the task performances. The phenomenon of differential perceptions based on cognitive styles may be too subtle for the simple tasks used in this experiment. Based on the results of this experiment, we can draw no conclusions about cognitive style as measured by this test and human performances in VR.

Conclusions

In this experiment we investigated conditions that theoretically could have made some difference in performance of simple tasks in VE. The finding that stereoscopic presentations are not significantly different than monoscopic presentations cannot be taken to mean that performance of the tasks is equivalent under these conditions. The rapid acquisition of skill was apparently sufficient to mask any small differences that may exist.

We have demonstrated that both stereoscopic presentation and head-coupling are of benefit when judging intermediate to short distances with a familiar target (as found in the **Distance Estimation** task). The stereoscopic display apparently provides some useful three dimensional cues for making less error in judging relatively short distances (one to three-plus meters). In addition, head-coupling apparently supports some use of parallax information across the same distances, which also decreases errors.

This research contributes to the importance of simulator sickness studies being part of the overall approach to VR investigations. There continue to be indications that certain configurations, task environments, and durations experienced in

VR systems have nauseogenic potential. As new technology is developed, these variables will require careful examination and experimentation in order to keep user comfort levels at a non-interfering level.

References

Bailey, J. (1994). Spatial knowledge acquisition. Unpublished doctoral dissertation.

Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? Presence, 1(3), 334-343.

Brown, J.J., & Baloh (1987). Persistent mal de barquement syndrome: A motion-induced subjective disorder of balance. American Journal of Otolaryngology, 8(4), 219-222.

Buckin, B.L. (1971). Field dependence and visual detection ability (Report No. 4137). Springfield, VA: National Technical Information Service.

Ekstrom, R.B., French, J.W., Harman, H.W., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests 1976. Educational Testing Service, Princeton, NJ.

Gordon, D.C., Spitzer, O., Shupak, A.,& Doweck, I. (1992). Survey of mal de barquement. British Medical Journal, 6826,544.

Hettinger, L.J., Berbaum, K.S., Kennedy, R.S., Dunlap, W.P., & Nolan, M.D. (1990). Vection and simulator sickness. Military Psychology, 2(3), 171-181.

Jacobs, R.S., Crooks, W.H., Crooks, J.R., Colburn, E., Fraser, R.E., Gorman, P.F., Madden, J.L., Furness, T.A., & Tice, S.E. (1993). Behavioral requirements for training and rehearsal in virtual environments (Final Report, Contract TR-9130-000-03-93). Illusion Engineering.

Kennedy, R.S., Fowlkes, J.E., & Lilienthal, M.G. (1993). Postural and performance changes following exposures to flight simulators. Aviation, Space, and Environmental Medicine, 64(10), 912-920.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. International Journal of Aviation Psychology, 3(3), 203-220.

Kirk, R. E. (1968). Experimental design: Procedures for the behavioral sciences. Belmont, CA: Wadworths.

Knerr, B.W., Goldberg, S.L., Lampton, D.R., Witmer, B.G., Bliss, J.P., Moshell, J.M., & Blau, B.S. (Spring, 1994). Research in the use of virtual environment technology to train dismounted soldiers. The Journal of Interactive Instruction Development, 6(4).

Kolasinski, E.M. (1995). Simulator sickness in virtual environments (ARI Technical Report 1027). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 295 861)

Koretz, J.F., & Handelmann, G.H. (1988). How the human eye focuses. Scientific American, 259(1), 92-99.

Lampton, D.R., Knerr, B.W., Goldberg, S.L., Bliss, J.P., Moshell, J.M., & Blau, B.S. (1994). The virtual environment performance assessment battery (VEPAB): Development and evaluation. Presence, 3(2).

Lipton, L. (1991). The crystal eyes handbook. San Rafael, CA: Stereographics.

Marascuilo, L.A., & Levin, J.R. (1970). Appropriate post hoc comparisons for interaction and nested interactions in analysis of variance designs: The elimination of type IV errors. American Educational Research Journal, 7(3), 397-421.

Murch, G.M. (1973). Visual and auditory perception. Indianapolis, IN: Bobs-Merril.

Murphy, T.P. (1993). Mal de barquement syndrome: A forgotten entity? Otolaryngology-Head and Neck Surgery, 109(1), 10-13.

Ono, H. (1991). Binocular visual directions of an object when seen as single or double. In D. Regan (Ed.), Binocular vision. Boca Raton, FL: CRC Press.

Patterson, R., & Martin, W.L. (1992). Human stereopsis. Human Factors, 34(6), 669-692.

Riccio, G.E., & Stoffregen, T.A. (1991). An ecological theory of motion sickness and postural instability. Ecological Psychology, 3(3), 195-240.

Rinalducci, E.J. (1993). Visual fidelity in the virtual environment. Unpublished manuscript. Orlando, FL: U.S. Army Research Institute for the Behavioral and Social Sciences.

Robinett, W., & Rolland, J.P. (1992). A computational model of the stereoscopic optics of a head-mounted display. Presence, 1(1), 45-62.

Rushton, S., & Wann, J. (1993, April). Problems in perception and action in virtual worlds. Paper presented at The Third Annual Conference on Virtual Reality, London, England.

Singer, M.J., Witmer, B., & Bailey, J.H. (1994). Development of "Presence" measures in virtual environments. Poster presented at the 38th Annual Meeting of the Human Factors and Ergonomics Society, Nashville, TN.

Sollenberger, R.L., & Milgram, P. (1993). Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. Human Factors, 35(3), 483-499.

Wann, J. (1993). What's wrong with your head mounted display? CyberEdge Journal, 5(Suppl. 17).

Ware, C., Arthur, K., & Booth, K.S. (1993). Fish tank virtual reality. Interchi '93. 24-29 April, 1993, 37-42.

Witmer, B.G., Bailey, J.H., & Knerr, B.W. (1995). Training dismounted soldiers in virtual environments: Route learning and transfer (ARI Technical Report 1022). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 292 900)

Witmer, B. & Singer, M.J. (1994). Measuring immersion in virtual environments (ARI Technical Report 1014). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 286 183)

Yeh, Y., & Silverstein, L.D. (1992). Spatial judgements with monoscopic and stereoscopic presentation of perspective displays. Human Factors, 34(5) 583-600.